



IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re the Application of

Grassy et al.

Application No. 09/359,181

Filed: July 22, 1999

# 21  
Plunkett  
3/24/02

For: "Method for the Provision, Identification and Description of Molecules Capable of Exhibiting a Desired Behavior More Particularly in the Pharmaceutical Sector and Molecules Obtained by Said Method"

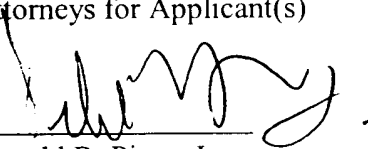
Hon. Commissioner of Patents  
and Trademarks  
Washington, D.C. 20231

SUBMISSION OF PRIORITY DOCUMENTS

Applicant hereby reaffirms the claim for the benefit of the filing dates of two corresponding applications: International Application No. FR/98/02909, filed on December 29, 1998 and Application No. 97/16706, filed on December 30, 1997. In accordance with the requirements of 35 USC §119 and 37 C.F.R. §1.55(a), we are filing herewith a certified copy of each of said applications upon which the claim for right of priority is made.

Respectfully submitted,

DANN, DORFMAN, HERRELL AND SKILLMAN  
A Professional Corporation  
Attorneys for Applicant(s)

By   
Donald R. Piper, Jr.  
Registration No. 29,337

Telephone (215) 563-4100  
Facsimile (215) 563-4044

The first part of the paper discusses the importance of understanding the local context in which a project is implemented. This includes a thorough understanding of the community's needs, values, and beliefs. It is essential to engage with the community from the very beginning, ensuring that their voices are heard and their input is valued. This process of community engagement is not a one-time event but a continuous one that evolves as the project progresses.

The second part of the paper explores the challenges that often arise in community-based projects. These challenges can range from a lack of resources to a lack of trust between the project team and the community. It is important to recognize these challenges early on and to develop strategies to address them. For example, building trust may require a long-term commitment and a willingness to listen to the community's concerns.

The third part of the paper discusses the importance of having a clear vision and a well-defined plan. This vision should be based on the community's needs and should be achievable within the available resources. A detailed plan should outline the steps that will be taken to implement the project, including a timeline and a budget. This plan should be shared with the community and should be subject to their input and approval.

The fourth part of the paper discusses the importance of monitoring and evaluation. This involves regularly assessing the progress of the project and making adjustments as needed. It is important to have a system in place to collect data and to analyze it in a way that provides useful information. This information can be used to identify areas where the project is working well and areas where it needs improvement.

The fifth part of the paper discusses the importance of sustainability. A project should be designed in a way that ensures it can continue to benefit the community long after the project team has left. This may involve training local people in the skills and knowledge needed to maintain the project. It may also involve establishing a local organization that will be responsible for the project's future.

In conclusion, the paper emphasizes the importance of a community-centered approach to project implementation. This approach involves listening to the community, engaging them in the process, and ensuring that the project is sustainable and meets their needs. It is a challenging but rewarding process that can lead to significant positive change in a community.



RECEIVED  
MAR 18 2002  
TECH CENTER 1000 2900

# BREVET D'INVENTION

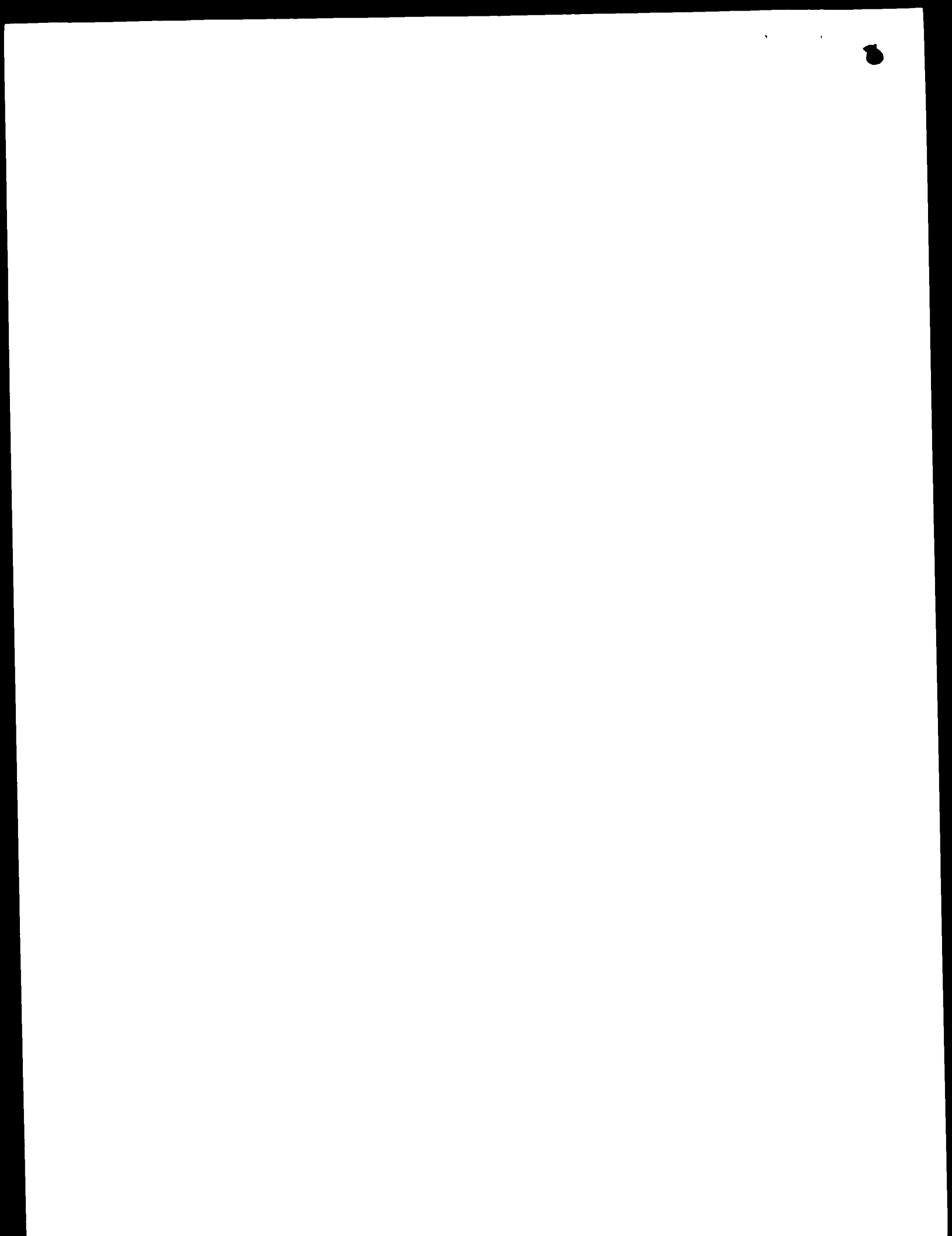
CERTIFICAT D'UTILITÉ - CERTIFICAT D'ADDITION

## COPIE OFFICIELLE

Le Directeur général de l'Institut national de la propriété industrielle certifie que le document ci-annexé est la copie certifiée conforme d'une demande de titre de propriété industrielle déposée à l'Institut.

En foi de quoi,

Martine PLANCHE

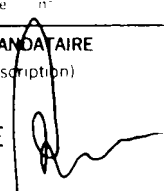



**REQUÊTE EN DÉLIVRANCE**

Confirmation d'un dépôt par télécopie

15 bis, rue de Saint-Petersbourg  
75800 Paris Cedex 08  
Téléphone : 01 53 04 53 04 Télécopie : 01 42 93 59 30

La requête est accompagnée de la somme de 100 francs en espèces ou par chèque.

Réserve à l'INPI DATE DE REMISE DES PIÈCES <b>30 DEC. 1997</b> N° D'ENREGISTREMENT NATIONAL <b>97 16706 -</b> DÉPARTEMENT DE DÉPÔT <b>π</b> DATE DE DÉPÔT <b>30 DEC. 1997</b>		1 NOM ET ADRESSE DU DEMANDEUR OU DU MANDATAIRE À QUI LA CORRESPONDANCE DOIT ÊTRE ADRESSÉE  <b>BREESE - MAJEROWICZ</b> <b>3, avenue de l'Opéra</b> <b>75001 PARIS</b>	
2 DEMANDE Nature du titre de propriété industrielle brevet d'invention demande internationale certificat d'utilité transformation d'une demande de brevet européen <div style="display: flex; align-items: center;"> <div style="border: 1px solid black; width: 15px; height: 15px; margin-right: 5px;"></div> <div style="text-align: center;">demande initiale</div> </div> brevet d'invention certificat d'utilité		n° du pouvoir permanent références du correspondant téléphone <b>S091B1118FR 01470367 77</b>	
Établissement du rapport de recherche différé immédiat Le demandeur, personne physique, requiert le paiement échelonné de la redevance oui non			
Titre de l'invention (200 caractères maximum)  <b>PROCEDE POUR PREVOIR, IDENTIFIER ET DECRIRE DES MOLECULES          SUSCEPTIBLES DE PRESENTER UN COMPORTEMENT RECHERCHE, NOTAMMENT          DANS LE DOMAINE DE LA PHARMACIE ET MOLECULES OBTENUES PAR CE PROCEDE</b>			
3 DEMANDEUR (S) SIREN Nom et prénoms (souligner le nom patronymique) ou dénomination  <b>SYNT:EM</b>		code APE-NAF Forme juridique  <b>S.A.</b>	
Nationalité (s) Française Adresse (s) complète (s)  <b>Parc Scientifique          Georges Besse          30000 Nîmes</b>		Pays  <b>FRANCE</b>	
4 INVENTEUR (S) Les inventeurs sont les demandeurs oui non Si la réponse est non, fournir une désignation séparée			
5 RÉDUCTION DU TAUX DES REDEVANCES requise pour la 1ère fois requise antérieurement au dépôt joindre copie de la décision d'admission			
6 DÉCLARATION DE PRIORITÉ OU REQUÊTE DU BÉNÉFICE DE LA DATE DE DÉPÔT D'UNE DEMANDE ANTÉRIEURE pays d'origine numéro date de dépôt nature de la demande			
7 DIVISIONS antérieures à la présente demande n° date n° date			
8 SIGNATURE DU DEMANDEUR OU DU MANDATAIRE (nom et qualité du signataire - n° d'inscription)  <b>Pierre BREESE</b> <b>921038</b>		SIGNATURE DU PREPOSÉ À LA RÉCEPTION SIGNATURE APRES ENREGISTREMENT DE LA DEMANDE À L'INPI   	



# BREVET D'INVENTION, CERTIFICAT D'UTILITÉ

DÉSIGNATION DE L'INVENTEUR

DIVISION ADMINISTRATIVE DES BREVETS

S91B1118FR

97/16706

TITRE DE L'INVENTION :

PROCEDE POUR PREVOIR, IDENTIFIER ET DECRIRE DES MOLECULES SUSCEPTIBLES DE PRESENTER UN COMPORTEMENT RECHERCHE, NOTAMMENT DANS LE DOMAINE DE LA PHARMACIE ET MOLECULES OBTENUES PAR CE PROCEDE.

LE(S) SOUSSIGNE(S)

BREESE-MAJEROWICZ  
3, avenue de l'Opéra  
75001 PARIS

DÉSIGNÉ(NT) EN TANT QU'INVENTEUR(S)

GRASSY Gérard  
23, rue du Pradas  
34470 PEROLS

KACZOREK Michel  
C/O  
Synt:em  
145, allée Charles Babbage  
30900 NIMES

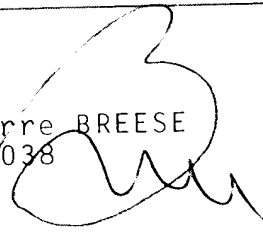
LAHANA Roger  
C/O  
Synt:em  
145, allée Charles Babbage  
3090 NIMES

YASRI Abdelaziz  
Synt:em  
145, allée Charles Babbage  
30900 NIMES

NOTE : Les auteurs de l'invention ont déclaré que l'invention est nouvelle, qu'elle est susceptible d'application industrielle et qu'elle est susceptible d'être brevetée.

Le 22 Mai 1998

Pierre BREESE  
921038



**DOCUMENT COMPORTANT DES MODIFICATIONS**

PAGE(S) DE LA DESCRIPTION OU DES REVENDECATIONS OU PLANCHE(S) DE DESSIN			R.M.*	DATE DE LA CORRESPONDANCE	TAMPON DATEUR DU CORRECTEUR
Modifiée(s)	Supprimée(s)	Ajoutée(s)			
19	-	-	-	20 mai 1998	- 4 JUIN 1998 B E P

[illegible]

1

PROCÉDÉ POUR PRÉVOIR, IDENTIFIER ET DÉCRIRE DES MOLÉCULES  
SUSCEPTIBLES DE PRÉSENTER UN COMPORTEMENT RECHERCHÉ,  
NOTAMMENT DANS LE DOMAINE DE LA PHARMACIE ET MOLECULES OBTENUES  
PAR CE PROCÉDÉ.

5

L'invention concerne la modélisation moléculaire.

Elle a pour objet un nouveau procédé utilisant un certain nombre de molécules de départ, notamment mais non seulement par chimie combinatoire virtuelle, pour prévoir, identifier et  
10 décrire des molécules susceptibles de présenter un comportement recherché, notamment dans le domaine de la pharmacie.

On sait que la recherche de nouvelles molécules actives, notamment dans le domaine de la pharmacie, nécessite la synthèse d'un très grand nombre de molécules qu'il faudra ensuite tester  
15 in vitro ou in vivo. Dans le meilleur des cas, seul un très petit nombre de ces molécules se révélera actif.

Pour tenter de rationaliser la recherche de nouvelles molécules actives on a eu l'idée de recourir à la modélisation moléculaire en utilisant des bases de données informatisées. Les  
20 résultats obtenus avec les techniques de ce type qui sont connues jusqu'à présent, ne sont pas toujours satisfaisants, en particulier en raison d'une définition insuffisante des paramètres et des critères d'activité.

La présente invention a pour but de remédier, au moins en  
25 partie, à ces inconvénients.

Ce but est atteint selon l'invention qui offre un procédé pour prévoir, identifier et décrire des molécules ayant un comportement recherché. Lequel procédé est basé sur la conception assistée par ordinateur et le criblage, également  
30 assisté par ordinateur, de banques combinatoires virtuelles. Ce procédé utilise différents descripteurs topologiques, de forme, chimiques, physiques etc. en combinaison avec une nouvelle analyse des trajectoires de dynamique moléculaire. Le terme de "comportement" signifie une "activité" au sens biologique ou  
35 pharmacologique, lorsque les molécules concernent des secteurs d'application pharmaceutique, ou une "propriété" au sens physico chimique, lorsque les molécules concernent des secteurs



d'applications non pharmaceutiques, par exemple des matériaux tels que des polymères.

Plus précisément, l'invention a pour objet un procédé assisté par ordinateur pour prévoir, identifier et décrire des molécules ayant un comportement recherché, notamment dans le domaine de la pharmacie, lequel procédé comprend essentiellement les étapes successives consistant à :

1) réaliser une base d'apprentissage à partir d'structures moléculaires voisines connues composées d'une part de structures moléculaires actives connues pour présenter l'activité recherchée et d'autre part de structures moléculaires inactives connues pour être dépourvues de cette activité ou pour présenter une activité recherchée faible, en utilisant des descripteurs appropriés ;

2) générer une série de structures moléculaires, préférentiellement de structures moléculaires voisines, notamment par explosion combinatoire, à partir de la base d'apprentissage ;

3) effectuer un criblage des molécules ainsi générées sur la base de l'enrichissement, en termes de diversité moléculaire, apporté par chaque molécule vis-à-vis des descripteurs choisis ;

4) soumettre les molécules ainsi retenues à des filtres statiques représentatifs des contraintes de variations structurelles, physico-chimiques et moléculaires auxquelles doivent répondre les molécules pour être actives et, éventuellement, synthétiser et tester les molécules sélectionnées ;

5) soumettre les molécules les plus prometteuses ainsi sélectionnées à un filtre dynamique représentatif des contraintes de variation conformationnelle que doivent respecter les molécules pour être actives ;

6) synthétiser et tester les molécules ainsi sélectionnées ;

7) si le résultat escompté n'est pas obtenu, ou s'il n'est obtenu que partiellement, répéter les étapes 3) à 6) en modifiant les filtres.

L'invention concerne également une molécule qui n'était pas antérieurement connue pour présenter un comportement recherché caractérisé en ce qu'elle occupe un espace conformationnel identique ou très voisin de l'espace conformationnel d'une molécule de référence au moins, la molécule de référence étant préalablement connue pour présenter le comportement recherché. La notion de "très voisin" s'apprécie en fonction de l'ensemble des espaces conformationnels correspondant aux molécules testées. Seront par exemple considérés comme "très voisins" les 10% d'espaces conformationnels les plus proches de l'espace conformationnel des molécules de références.

Lorsque le procédé s'appuie sur une pluralité de molécules de références, on considérera comme "très voisins" les espaces conformationnels des molécules de référence.

Pour réaliser une base d'apprentissage, on utilise des descripteurs variés qui peuvent être toutes sortes de propriétés quantitatives et/ou semi-quantitatives. Par exemple, comme cela est connu de l'homme du métier, à partir de la représentation graphique d'une molécule, en d'autres termes une représentation en deux dimensions, ou représentation 2D, d'un composé chimique, on peut déduire un ensemble de valeurs numériques appelées "descripteurs topologiques". On peut en outre utiliser des descripteurs reflétant certaines propriétés physico-chimiques comme par exemple le caractère lipophile, ou lipophilie, qui est exprimé en tant que  $\log P$ ,  $P$  étant le coefficient de partition du composé considéré entre l'eau et le n-octanol, ou la réfractivité molaire. On peut en outre utiliser des descripteurs numériques représentant des formes moléculaires.

Les techniques classiquement utilisées dans les correlations quantitatives de structure-activité (Quantitative Structure Activity Relationships : QSAR) sont basées sur l'hypothèse que si une molécule présente un comportement biologique donné, toutes les informations nécessaires pour la caractériser se fondent fortement sur sa structure, c'est-à-dire ses atomes, ses liaisons et ses formes. Contrairement au paradigme QSAR linéaire classique dans lequel l'activité

biologique peut être exprimée sous la forme d'une combinaison linéaire de descripteurs pertinents, la Demanderesse utilise un paradigme de représentation d'informations ou mappage variable, non linéaire, dans lequel l'activité est une fonction non  
5 linéaire de descripteurs structuraux, topologiques et moléculaires.

Selon l'invention, on analyse un ensemble de descripteurs pour les composés actifs, en opposition aux composés inactifs, ce qui permet la définition de filtres pour faire la différence  
10 entre les deux classes de molécules [G. Grassy & al., J. Mol. Graphics, vol. 13, page 356 (1995)].

Si l'échantillon d'apprentissage correspond à des tests *in vivo*, le jeu de filtres décrit les conditions d'activité *in vivo*. Dans le cas des tests *in vitro*, le jeu de filtres décrit  
15 les conditions d'activité *in vitro*.

Un filtre est défini par la plage de variation d'un descripteur donné pour tous les composés actifs connus lorsqu'on la compare à la plage de variation du même descripteur pour tous les composés inactifs connus.

20 Si la plage "active" recouvre totalement la plage "inactive", le filtre ne sert à rien, ce qui revient à dire que la variation de ce descripteur particulier n'a pas de relation avec l'activité biologique.

Si la plage "active" ne recouvre pas totalement la plage  
25 "inactive", toute molécule présentant pour ce descripteur une valeur qui tombe dans une zone de la plage "active" qui ne recouvre pas la zone "inactive" aura une forte probabilité d'être active aussi.

Les aptitudes à la prédiction de ce type d'approche sont  
30 améliorées par l'utilisation de plusieurs filtres, préférentiellement 10 à 30, chacun représentant une manière différente de quantifier les propriétés structurales et physico-chimiques d'une molécule.

Ces filtres, une fois établis, peuvent être utilisés pour  
35 cribler des composés virtuels générés pendant une explosion combinatoire, c'est-à-dire tous les composés obtenus par exemple en branchant une liste de substituants sur des parties

prédéterminées d'un noyau structurel appelé synthon. Les composés les plus prometteurs, à savoir ceux qui satisfont à toutes les contraintes définies par les filtres, sont ensuite synthétisés par synthèse chimique ou par génie génétique, ou encore par tout autre moyen, et testés dans des essais biologiques.

Les molécules inactives que l'on avait prédites actives contiennent des caractéristiques inexplorées qui sont déterminantes pour l'activité. De telles propriétés sont utilisées pour la définition de filtres additionnels et, après quelques répétitions, le processus de criblage virtuel peut être utilisé pour l'identification exacte de composés présentant les propriétés recherchées.

En pratique, à partir des structures de départ dont l'activité peut être mesurée *in vitro* et/ou *in vivo*, on réalise une "base d'apprentissage" en déterminant les différents paramètres physico-chimiques, structuraux et moléculaires permettant de décrire les molécules actives et les molécules inactives.

Ces paramètres sont représentés sous forme de plages de variations, ou "contraintes", en fonction des classes d'activité. Un jeu ou ensemble de contraintes définit un filtre.

La qualité des descripteurs choisis est contrôlée en termes de variabilité, ainsi que le choix des descripteurs à utiliser et l'intercorrélation entre les descripteurs retenus.

Ceci peut être réalisé en utilisant un logiciel conçu par R. Lahana, le logiciel ANODA (ANalysis Of DATA) qui associe à une chimiothèque une "carte de visite" simple en termes de descripteurs de base et d'analyse en composantes principales. Différentes techniques de sélection des descripteurs telles qu'analyses de variabilité, d'intercorrélation, de représentativité (régressions, réseaux de neurones, algorithmes génétiques) sont mises en oeuvre.

Une explosion combinatoire est générée à partir de la base d'apprentissage et en prenant en compte les conditions de variation utiles pour chaque position variable. Pour ce faire,

il est possible d'utiliser des bases de données de substituants définis par leur structure et les descripteurs qui en découlent.

5 L'explosion combinatoire peut être générée en utilisant du logiciel LEGION de la société TRIPOS ou un autre logiciel conçu par R. Lahana, le logiciel COMBEX qui crée une telle explosion à partir d'un répartiteur ("scaffold") ou d'une séquence consensus, et d'une liste de contraintes sur chacun des points de substitution. Un langage de commande de type SQL (Standard query language) permet de combiner à volonté toutes sortes de  
10 conditions de sélection des substituants.

Pour chaque combinaison générée, il est possible de vérifier l'enrichissement en terme de diversité moléculaire apportée par la molécule créée vis-à-vis des descripteurs choisis. Si cet enrichissement est suffisant, la molécule est  
15 conservée, sinon elle est rejetée.

L'enrichissement en terme de diversité moléculaire peut être vérifié au moyen d'un logiciel conçu par R. Lahana, le logiciel DIVERSE qui permet d'effectuer une évaluation quantitative de la diversité moléculaire présentée par une  
20 chimiothèque quelconque, même si les molécules qui la composent sont hautement flexibles et si cette chimiothèque est purement virtuelle. Ce logiciel permet de comparer des banques en terme de diversité, de caractériser des "trous" de diversité dans une banque, de concevoir rationnellement des banques minimales ayant  
25 une diversité maximale. Les techniques de sélection peuvent être par exemple la classification hiérarchique (clusters), des partitions, la conception uniforme (uniform design), le tirage aléatoire, etc.

Les filtres "statiques", c'est-à-dire ceux qui ne  
30 dépendent pas de la variabilité conformationnelle des molécules, décrivent les intervalles des valeurs structurales, physico-chimiques, moléculaires, auxquels les molécules doivent correspondre afin d'être actives. En utilisant ces filtres statiques, on déduit les molécules ayant le plus de chances  
35 d'être actives. Ces molécules sont éventuellement synthétisées et testées pour l'activité recherchée.

Dans cette étape, on peut utiliser un logiciel conçu par G. Grassy, le logiciel VARIMAP qui établit des filtres statiques à partir de descripteurs obtenus par TSAR qui est un logiciel commercialisé par Oxford Molecular Group Oxford (Royaume-Uni) et dont les auteurs initiaux sont G. Grassy et R. Lahana.

Les candidats les plus prometteurs sélectionnés au moyen des filtres statiques sont ensuite soumis à un filtre dynamique, c'est-à-dire représentatif des contraintes de variation conformationnelle que doivent respecter les molécules générées pour avoir des chances d'être actives, par comparaison avec celles de la base d'apprentissage.

Ce filtrage dynamique peut être effectuée au moyen d'un logiciel nouvellement conçu par G. Grassy, le logiciel MULTIDYN qui permet, à partir de trajectoires de dynamique moléculaire, de caractériser les espaces conformationnels de molécules quelconques. Les conformations bioactives des molécules en question sont ainsi mises en évidence.

Les molécules retenues sont ensuite synthétisées et testées.

Si l'activité biologique trouvée ne répond pas ou ne répond que partiellement aux critères attendus, les étapes de vérification de l'enrichissement en terme de diversité moléculaire, de filtrage statique, de filtrage dynamique, de synthèse et d'essais sont répétées.

Si on trouve que des molécules répondent aux critères biologiques fixés, le but est atteint.

Le procédé selon l'invention permet donc de définir de nouvelles molécules actives par un processus entièrement rationnel n'utilisant à aucun moment d'hypothèse sur leur mécanisme d'action possible ou sur leur récepteur éventuel.

Le procédé selon l'invention a été mis en oeuvre avec succès dans des domaines de pharmaco-modulation de molécules afin d'améliorer les performances de ces molécules *in vitro* et surtout *in vivo*.

L'exemple détaillé de mise en oeuvre qui suit sert à illustrer et mieux expliquer l'invention.

#### EXEMPLE

### Obtention de nouveaux composés immuno-modulateurs

On a récemment montré que le peptide 2702.75-84 (peptide dérivé du HLA-B2702, acides aminés 75 à 84) prolonge, chez la souris, la survie après une allo-greffe du

5 coeur. (Transplantation, vol. 59, page 455 (1995)).

L'administration du peptide 2702.75-84 à raison de 80 mg/kg/jour pendant 10 jours après la transplantation de coeurs B6 chez des receveurs CBA prolonge la survie à l'allo-greffe de coeur jusqu'à 11,4 ± 2,6 (n=8) jours comparés à 8,2 ± 1,2

10 jours chez les animaux témoins non traités ( $p < 0,01$ ). Aucun effet sur la survie à la greffe n'a été observé après une thérapie avec le peptide 2702.75-84 à des doses inférieures.

L'activité *in vivo* de plusieurs peptides (n=19) dérivés du peptide 2702.75-84 et d'autres séquences MHC/HLA a été évaluée

15 de façon similaire. Les peptides testés comprenaient des peptides de types acides aminés D et L. Certains peptides différaient du peptide 2702.75-84 sur jusqu'à six positions d'acides aminés, d'autres avaient une séquence inversée.

Tous ces peptides dont les séquences sont représentées

20 dans le tableau I qui suit, ont été synthétisés en utilisant la chimie F-moc/tBu puis testés.

Les études de réponse aux doses sur des souris ayant reçu une allo-greffe de coeur ont été effectuées en utilisant des peptides sous la forme acétate, purifiés à plus de 90% par

25 chromatographie liquide haute performance (HPLC).

La transplantation cardiaque hétérotrope abdominale a été effectuée comme décrit précédemment J. Thorac. Cardiovasc. Surg., vol. 7, page 225 (1969).

Des souris CBA receveuses de coeurs C57Bl/6 ont été

30 traitées quotidiennement avec différentes doses de peptide après la transplantation d'organe. Les peptides étaient dissous dans du DMSO et dilués dans du PBS (concentration finale en DMSO 10%) avant l'administration intrapéritonéale. Les animaux ont été traités à partir du jour de la transplantation jusqu'au 9ème

35 jour. La survie à la greffe a été surveillée quotidiennement par palpation directe et le rejet a été défini comme la fin de la contractilité cardiaque palpable. La signification statistique de

la prolongation de survie à l'allo-greffe du coeur a été calculée en utilisant le test de Mann-Whitney.

5 La thérapie des receveurs de coeur de souris allo-greffe avec certains de ces peptides ( $n=9$  ; en gras dans le tableau I qui suit) a permis une prolongation significative de la survie à la greffe, alors que les autres n'avaient pas d'effets significatifs.

Les structures des peptides utilisés et les résultats obtenus sont présentés dans le tableau I qui suit.

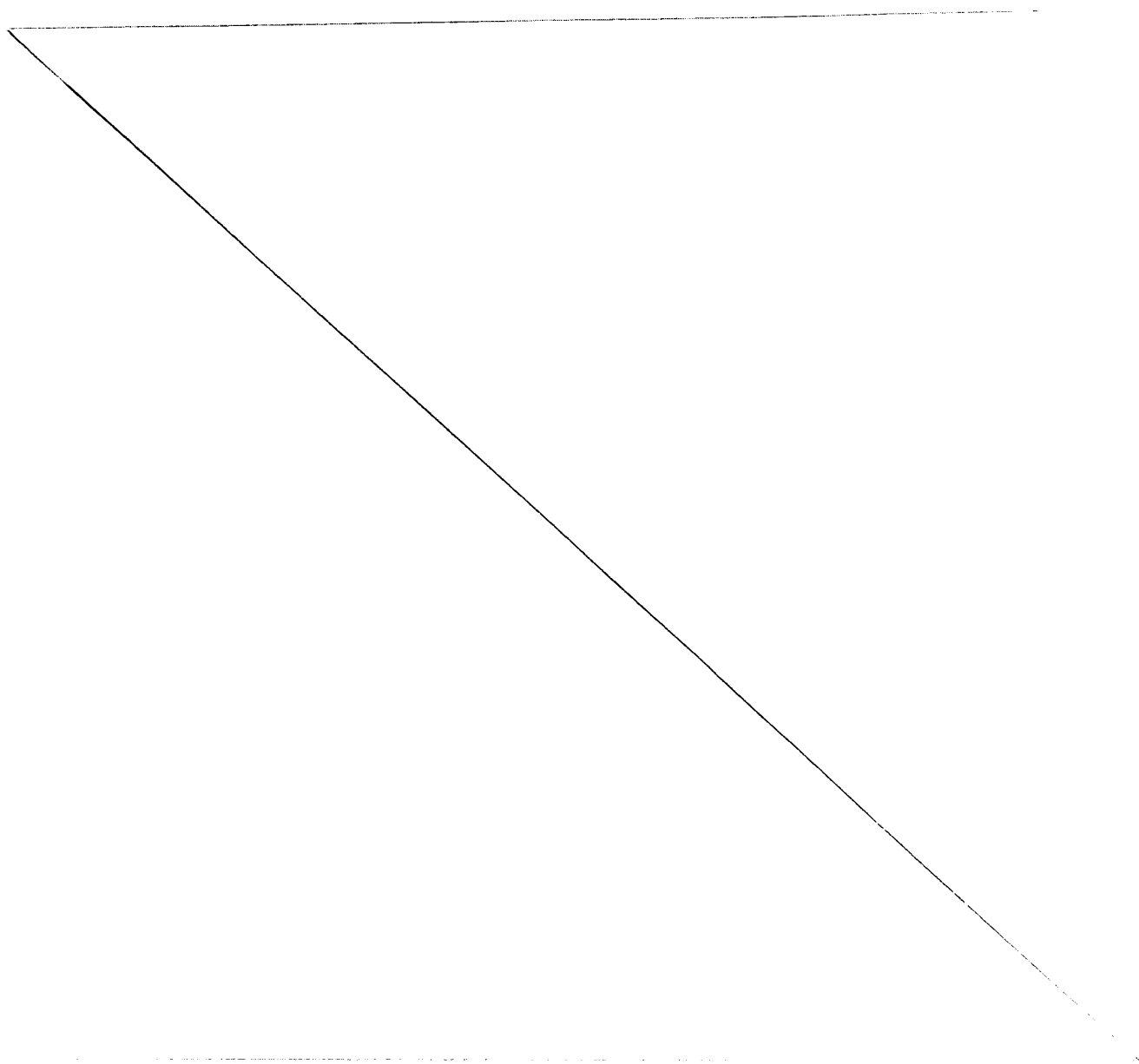




TABLEAU I

Séquence peptidique	HLA/MHC	MSTSD
Non traite		7,5 1,1
RENLRIALPY	B2702	11,4 2,6
YRLAIRLNEF	-	12,1 2,8
renlrrialry	-	11,4 4,1
yrlairlner	-	13,2 2,7
RVNLFIALEY	-	11,5 0,5
YRLAIRLNVF	-	12,5 1,6
rvnlrialry	-	13,1 3,9
yrlairlnvr	-	12,2 2,9
NLRIALRYYW	-	11,6 1,3
RVNLFITALFY	Kk	8,5 0,7
FVDLRTLLEY	Dk	7,0 0,5
FVDKETLLGY	Kb	7,8 1,0
FVSLFNLLGY	Db	8,0 0,5
FESLRLLRGY	07	7,5 0,7
FEDLRTLLEY	B2705	7,7 1,2
ENLRIALR	-	8,5 0,7
renlpialry	-	9,5 2,4
RVNLRTLRRY	E	8,0 0,5
RMNLQTLRGY	G	7,5 0,7

5 L'ensemble des 19 peptides a été utilisé comme base d'apprentissage pour définir la stratégie de conception rationnelle.

Bien que cet ensemble de départ soit très petit, il a permis, comme le démontre la distinction initiale entre les composés actifs et les composés inactifs, une définition efficace des contraintes.

10 Initialement, 27 descripteurs indépendants de la conformation ont été calculés pour chaque peptide (voir le Tableau II ci après).

Les descripteurs physico-chimiques et topologiques ont été générés par le logiciel TSAR V2.31 (Oxford Molecular Group, Oxford, Royaume-Uni).

5 Ces 27 descripteurs comprenaient le moment dipolaire, calculé sur la base d'une conformation totalement développée de chaque peptide. L'analyse statistique a montré que 14 descripteurs étaient intercorrélés. Par conséquent, ils n'étaient pas utiles pour la définition de contraintes qui différencient entre les peptides actifs et les inactifs. Les 13  
10 autres descripteurs, indépendants de la conformation, ont été utilisés pour construire un filtre statique pour le criblage d'une bibliothèque combinatoire virtuelle comme décrit plus loin.

TABLEAU II

15

Les 27 descripteurs topologiques  
et physicochimiques utilisés

Propriété	Nature	Inclus/exclu
<i>Masse molaire</i>	<i>Physique</i>	<i>Exclu</i>
<i>Volume ellipsoïdal</i>	<i>Forme</i>	<i>Inclus</i>
<i>Volume moléculaire</i>	<i>Forme</i>	<i>Inclus</i>
<i>Réfractivité molaire</i>	<i>Topologique</i>	<i>Inclus</i>
<i>Lipophilie (LogP)</i>	<i>Topologique</i>	<i>Inclus</i>
<i>Kappa 1</i>	<i>Topologique</i>	<i>Exclu</i>
<i>Kappa 2</i>	<i>Topologique</i>	<i>Exclu</i>
<i>Kappa 3</i>	<i>Topologique</i>	<i>Exclu</i>
<i>Kappa Alpha 1</i>	<i>Topologique</i>	<i>Exclu</i>
<i>Kappa Alpha 2</i>	<i>Topologique</i>	<i>Inclus</i>
<i>Kappa Alpha 3</i>	<i>Topologique</i>	<i>Exclu</i>
<i>Flexibilité</i>	<i>Topologique</i>	<i>Inclus</i>
<i>Kier Chi V4</i>	<i>Topologique</i>	<i>Inclus</i>
<i>Indice Randic</i>	<i>Topologique</i>	<i>Exclu</i>
<i>Indice Balaban</i>	<i>Topologique</i>	<i>Inclus</i>
<i>Indice Wiener</i>	<i>Topologique</i>	<i>Exclu</i>
<i>Somme état E</i>	<i>Physique</i>	<i>Exclu</i>

Moment dipolaire	Physique	Inclus
Nombre d'atomes de C	Chimique	Exclu
Nombre d'atomes de O	Chimique	Inclus
Nombre d'atomes de N	Chimique	Inclus
Nombre d'atomes de H	Chimique	Exclu
Nombre total d'atomes	Chimique	Exclu
Nombre de groupes méthyle	Chimique	Exclu
Nombre de groupes éthyle	Chimique	Inclus
Nombre de groupes amino	Chimique	Exclu
Nombre de groupes hydroxy	Chimique	Inclus

Les descripteurs topologiques listés ci-dessus ont été utilisés pour calculer les filtres statiques. Treize descripteurs (en écriture droite), ont fourni des informations sur les caractéristiques des peptides ayant une activité immuno-

5 modulatrice et ont été utilisés pour définir des contraintes pour cribler une bibliothèque combinatoire virtuelle. Les quatorze autres descripteurs se sont révélés être intercorrélés et ont été exclus de l'analyse.

10 En utilisant le programme COMBEX (Synt:em, Nîmes, France), on a généré une explosion combinatoire basée sur une séquence consensus RXXRXRXXXY, dérivée de l'ensemble d'apprentissage, après alignement de toutes les séquences actives et inactives. Cette séquence a laissé 7 positions, les positions représentées

15 par "X", pour effectuer des mutations afin de créer la bibliothèque.

Toutes les molécules ont été générées en utilisant la convention SMILES puis converties en une structure 3D en utilisant le logiciel CORINA (Oxford Molecular Group, Oxford,

20 Royaume-Uni).

Initialement, des acides aminés tant naturels que non naturels étaient compris dans la base de données de substituants. Tous les acides aminés étaient décrits en termes de propriétés physico chimiques (lipophilie, basicité,

25 aromaticité, etc.) et également par des descripteurs topologiques (analyse de Kier, indice Balaban, etc.). En

utilisant 35 acides aminés, cela conduisait à 357 combinaisons, soit 64 milliards de composés, ce qui était encore trop important pour la capacité de l'ordinateur utilisé.

5 Afin de diminuer ce nombre, quelques données supplémentaires telles que la distribution de la lipophilie, ont été prises en compte. En fait, la structure du composé "conducteur" (2702.75-84) présentait deux domaines lipophiles séparés par des résidus hydrophiles. Ceci suggérait qu'une telle distribution était importante pour l'activité. Afin de respecter  
10 cette distribution, la Demanderesse a décidé d'utiliser, pour chacune des 7 positions, la liste suivante d'acides aminés : V, I, T, W, L, nL (nL = norleucine). Ainsi, le nombre des composés de la bibliothèque a été ramené à 67 combinaisons, soit 279 936 composés.

15 Pour cribler la bibliothèque combinatoire virtuelle, on a calculé "au vol", pour chaque structure générée par le logiciel COMBEX, l'ensemble correspondant de propriétés. Ces propriétés ont été analysées en utilisant les filtres statiques et les filtres dynamiques prédéfinis, basés sur les contraintes et on  
20 n'a retenu que les composés satisfaisant à toutes les contraintes.

Les filtres statiques, obtenus par les plages de variation des descripteurs retenus étaient basés sur l'ensemble de contraintes indépendantes de la conformation défini en utilisant  
25 l'ensemble d'apprentissage de composés actifs et inactifs. Ils ont été conçus en utilisant le logiciel VARIMAP (Synt:em, Nîmes, France).

En criblant la bibliothèque de 279 936 composés avec ces filtres statiques, on a identifié 26 peptides qui satisfaisaient  
30 à toutes les contraintes. Parmi ceux-ci les peptides ont été étudiés du point de vue de leurs espaces conformationnels (filtre dynamique) et en particulier 5, appelés respectivement RDP1257, RDP1258, FDP1259, RDP1271 et RDP1277 ont été étudiés.

Le caractère flexible des peptides a été analysé en  
35 utilisant des simulations de dynamique moléculaire (MD).

Les simulations MD des peptides solvatés avec des conditions périodiques ont été effectuées en utilisant le

logiciel AMBER 4.1 (Oxford Molecular Group, Oxford, Royaume-Uni). Il utilisait une durée de 1015 ps pour chaque peptide solvaté. La constante diélectrique été réglée sur l'unité. La température du système a été initialement portée progressivement de 10 à 300 K, sur une période de 15 ps. Pendant la simulation, on a maintenu une température constante de  $300 \pm 10$  K en couplant a un bain extérieur avec une durée de relaxation de 0,1 ps. Le "pas" temporel choisi était de 1 fs. Le temps de calcul était approximativement de 0,5 h par ps. On a utilisé une coupure "basée sur les résidus" de 10 Å pour toutes les interactions non liées. La liste des paires non liées était remise à jour toutes les 10 fs et les coordonnées étaient recueillies chaque ps pendant les trajectoires, ce qui a donné un ensemble de 1015 conformations pour chaque trajectoire. Dans toutes les trajectoires, aucune contrainte n'a été appliquée aux atomes et aucun "terme parasite" n'a été utilisé dans l'expression d'énergie.

Chaque conformation a été représentée par un descripteur de forme appelé le vecteur d'autocorrélation 3D ci-après "3D-ACV" [Eur. J. Med. Chem.- Chim. Theor., vol. 19, page 61 (1984)]. Un ensemble de 3D-ACV a été calculé pour chaque série de conformations obtenues par dynamique moléculaire (MD) puis traité en utilisant les statistiques à plusieurs variables [Trends in QSAR and Molecular modelling 92 ESCOM Publishers, page 216 (1993)]. Ceci a été réalisé en trois étapes principales (i à iii):

(i) Pour une conformation donnée de la molécule étudiée, ici un peptide, le descripteur 3D-ACV correspondant a été calculé comme suit. Les distances entre toutes les paires d'atomes ont été calculées. La distribution de ces distances était un vecteur dont chaque "boîte" (bin) était la somme de paires d'atomes dans une gamme spécifique de séparation interatomique, c'est-à-dire où deux atomes donnés étaient séparés par une distance comprise entre  $(r-1)$  et  $r$  Å. Dans ce travail, le pas était égal à 1 Å. Evidemment, la moindre modification de la conformation de ladite molécule aboutit à un changement de la distribution des distances interatomiques, donc

à une modification du 3D-ACV. Par conséquent, ce descripteur est l'un des plus efficaces pour représenter les formes conformationnelles des molécules.

5 (ii) Une trajectoire MD est l'ensemble des conformations adoptées par une molécule donnée pendant la simulation MD. Pour chaque conformation, le 3D-ACV correspondant a été calculé "au vol" et mémorisé. Ce 3D-ACV multiple, fonction du temps est un descripteur du comportement dynamique de ladite molécule.

10 (iii) Pour la comparaison des 3D-ACV multiples représentant les trajectoires de l'ensemble de molécules analysé, on a appliqué une analyse de composants principaux (PCA) à chacun de ces 3D-ACV multiples. Cette transformation a réduit les dimensions de l'ensemble de données à un nombre plus faible (dans le cas présent, un espace 2D) et a aussi permis la  
15 projection de toutes les trajectoires de toutes les molécules sur un plan. Dans cet espace réduit, chaque molécule est représentée par un ensemble de points, c'est-à-dire ses conformations pendant toute la durée de la simulation MD, ce qui représente son espace conformationnel. Les molécules pouvaient  
20 ensuite être comparées l'une à l'autre en termes d'espaces conformationnels. Les analyses des trajectoires et des espaces conformationnels ont été calculés au moyen du logiciel MULTIDYN (Synt:em, Nîmes, France).

Par utilisation d'un filtre dynamique pour cribler les  
25 peptides en termes d'espaces conformationnels, il est apparu que 4 de ces 5 peptides occupaient le même espace conformationnel ou un espace très similaire mais que l'un, le RDP1277, était différent de ce point de vue.

30 Ces 5 peptides ont été synthétisés et testés dans des essais biologiques de survie à l'allo-greffe du coeur.

A une dose de 10 mg/kg/jour, tous ces peptides, sauf le RDP1277, étaient actifs *in vivo*. Aucune prolongation significative de survie à la greffe n'a été observée après la  
35 thérapie avec le RDP1277 (MST = 9,0 ± 1,4). En revanche, la thérapie avec tous les autres peptides a permis une prolongation significative de la survie à la greffe, allant de 11 à 13 jours.

Des essais supplémentaires avec le peptide RDP1258 à des doses de 1 à 10 mg/kg/jour ont révélé une efficacité accrue de la thérapie aux doses inférieures. A une dose de 1 mg/kg/jour, 30% des souris ayant reçu une allo-greffe de coeur et traitées  
5 par le RDP1258 on survécu pendant plus de 100 jours, alors que le peptide de référence, 2705.75-84, prolongeait la survie à l'allo-greffe de coeur après une thérapie à raison de 80 mg/kg/jour, tandis qu'aucun effet n'était observé aux doses inférieures. De plus, aucune survie de longue durée à la greffe  
10 n'a été induite par une thérapie utilisant ce peptide de référence.

La stratégie selon l'invention a donc bien permis de concevoir de façon rationnelle plusieurs composés bioactifs dont l'activité s'est révélée supérieure à celle des molécules  
15 actives de la base d'apprentissage.

## REVENDICATIONS

1. Procédé assisté par ordinateur pour  
prévoir, identifier et décrire des molécules ayant un  
5 comportement recherché, notamment dans le domaine de la  
pharmacie, mettant en oeuvre une étape de modélisation  
moléculaire, une étape de constitution d'une banque combinatoire  
et une étape de sélection des molécules potentiellement  
intéressantes, caractérisé en ce qu'il comporte une étape de  
10 filtration des molécules candidates par un filtre dynamique  
représentatif des contraintes de variations conformationnelles  
que doivent respecter les molécules pour présenter ladite  
activité ou par un filtre statique représentatif d'une pluralité  
de descripteurs.

15 2. Procédé assisté par ordinateur pour  
prévoir, identifier et décrire des molécules ayant un  
comportement recherché, notamment dans le domaine de la  
pharmacie, selon la revendication 1 caractérisé en ce qu'il  
comporte une étape préalable de filtrage statique pour réduire  
20 le nombre de molécules candidates et pour une étape ultérieure  
de filtration dynamique.

3. Procédé assisté par ordinateur pour prévoir,  
identifier et décrire des molécules ayant un comportement  
recherché, notamment dans le domaine de la pharmacie, selon la  
25 revendication 1, caractérisé en ce qu'il comprend  
essentiellement les étapes successives consistant à :

1) réaliser une base d'apprentissage à partir d'une  
famille de structures moléculaires connues composée d'une part  
de structures moléculaires actives présentant l'activité  
30 recherchée et d'autre part de structures moléculaires inactives  
connues dépourvues de cette activité, ou comme présentant une  
activité faible, en utilisant des descripteurs appropriés ;

2) générer une explosion combinatoire de molécules à  
partir de la base d'apprentissage ;

35 3) effectuer un criblage des molécules ainsi  
générées sur la base de l'enrichissement, en termes de diversité



moléculaire, apporté par chaque molécule vis-à-vis des descripteurs choisis ;

4) soumettre les molécules ainsi retenues à des filtres statiques représentatifs des contraintes de variations structurelles, physico-chimiques et moléculaires auxquelles doivent répondre les molécules pour être actives et, éventuellement, synthétiser et tester les molécules sélectionnées;

5) soumettre les molécules les plus prometteuses ainsi sélectionnées à un filtre dynamique représentatif des contraintes de variation conformationnelle que doivent respecter les molécules pour être actives ;

6) synthétiser et tester les molécules ainsi sélectionnées.

4. Procédé assisté par ordinateur pour prévoir, identifier et décrire des molécules ayant un comportement recherché, notamment dans le domaine de la pharmacie, selon la revendication 3, caractérisé en ce que, si le résultat escompté n'est pas obtenu, ou s'il n'est obtenu que partiellement, les étapes 3) à 6) sont répétées en modifiant les filtres.

5. Procédé assisté par ordinateur pour prévoir, identifier et décrire des molécules ayant un comportement recherché, notamment dans le domaine de la pharmacie, selon la revendication 3 ou 4, caractérisé en ce que le filtrage dynamique peut être effectuée au moyen d'un logiciel qui permet, à partir de trajectoires de dynamique moléculaire, de caractériser les espaces conformationnels de molécules quelconques.

6. Procédé assisté par ordinateur pour prévoir, identifier et décrire des molécules ayant un comportement recherché, notamment dans le domaine de la pharmacie, selon l'une quelconque des revendications précédentes, caractérisé en ce que, le filtrage dynamique met en oeuvre le descripteur de forme constitué par le vecteur d'autocorrélation 3D (3D-ACV).

7. Procédé assisté par ordinateur pour prévoir, identifier et décrire des molécules ayant un comportement recherché, notamment dans le domaine de la pharmacie, selon la

revendication 6, caractérisé en ce que le filtrage dynamique consiste à calculer un ensemble de 3D-ACV pour chaque série de conformations obtenues par dynamique moléculaire (MD) puis traité en utilisant les statistiques à plusieurs variables en

5 trois étapes principales (i à iii):

(i) Pour une conformation donnée de la molécule étudiée, on calcule le descripteur 3D-ACV correspondant par détermination des distances entre toutes les paires d'atomes et de la distribution de ces distances formant un vecteur dont chaque

10 "boîte" (bin) est la somme de paires d'atomes dans une gamme spécifique de séparation interatomique,

(ii) On calcule et on mémorise pour chaque conformation, le 3D-ACV multiple correspondant "au vol" ;

(iii) On procède à une analyse de composants

15 principaux (PCA) à chacun de ces 3D-ACV multiples.

9. Procédé assisté par ordinateur pour prévoir, identifier et décrire des molécules ayant un comportement recherché, notamment dans le domaine de la pharmacie, selon la revendication 5, caractérisé en ce que le filtre statique met en

20 oeuvre des descripteurs physico-chimiques et topologiques dont une partie au moins est choisie parmi les descripteurs cités dans le tableau II.

9. Molécule ayant un comportement recherché, notamment dans le domaine de la pharmacie, conçue par le procédé

25 conforme à l'une au moins des revendications précédentes.

10. Molécule qui n'était pas antérieurement connue pour présenter un comportement recherché, caractérisé en ce qu'elle occupe un espace conformationnel identique ou très voisin de l'espace conformationnel d'une molécule de référence

30 au moins, la molécule de référence étant préalablement connue pour présenter le comportement recherché.

11. Molécule qui n'était pas antérieurement connue pour présenter un comportement recherché, caractérisé en ce qu'elle présente des plages de variation des descripteurs

35 statiques identique ou très voisin des plages de variation des descripteurs statiques d'une molécule de référence au moins, la

molécule de référence étant préalablement connue pour présenter le comportement recherché.

12. Molécule qui n'était pas antérieurement connue pour présenter un comportement recherché, selon la revendication 5 10 et 11 caractérisé en ce qu'elle occupe un espace conformationnel identique ou très voisin de l'espace conformationnel d'une molécule de référence au moins, la molécule de référence étant préalablement connue pour présenter le comportement recherché et en ce qu'elle présente des plages 10 de variation des descripteurs statiques identique ou très voisin des plages de variation des descripteurs statiques d'une molécule de référence au moins, la molécule de référence étant préalablement connue pour présenter le comportement recherché.

revendication 6, caractérisé en ce que le filtrage dynamique consiste à calculer un ensemble de 3D-ACV pour chaque série de conformations obtenues par dynamique moléculaire (MD) puis traité en utilisant les statistiques à plusieurs variables en  
5 trois étapes principales (i à iii):

(i) Pour une conformation donnée de la molécule étudiée, on calcule le descripteur 3D-ACV correspondant par détermination des distances entre toutes les paires d'atomes et de la distribution de ces distances formant un vecteur dont chaque  
10 "boîte" (bin) est la somme de paires d'atomes dans une gamme spécifique de séparation interatomique,

(ii) On calcule et on mémorise pour chaque conformation, le 3D-ACV multiple correspondant "au vol" ;

(iii) On procède à une analyse de composants  
15 principaux (PCA) à chacun de ces 3D-ACV multiples.

8. Procédé assisté par ordinateur pour prévoir, identifier et décrire des molécules ayant un comportement recherché, notamment dans le domaine de la pharmacie, selon la revendication 5, caractérisé en ce que le filtre statique met en  
20 oeuvre des descripteurs physico-chimiques et topologiques dont une partie au moins est choisie parmi les descripteurs cités dans le tableau II.

9. Molécule ayant un comportement recherché, notamment dans le domaine de la pharmacie, conçue par le procédé  
25 conforme à l'une au moins des revendications précédentes.

10. Molécule qui n'était pas antérieurement connue pour présenter un comportement recherché, caractérisé en ce qu'elle occupe un espace conformationnel identique ou très voisin de l'espace conformationnel d'une molécule de référence  
30 au moins, la molécule de référence étant préalablement connue pour présenter le comportement recherché.

11. Molécule qui n'était pas antérieurement connue pour présenter un comportement recherché, caractérisé en ce qu'elle présente des plages de variation des descripteurs  
35 statistiques identique ou très voisin des plages de variation des descripteurs statistiques d'une molécule de référence au moins, la